DOCUMENT RESUME

ED 035 589 24 SP 003 447

AUTHOR	Good, Thomas L.; And Others
TITLE	Teacher Assessment of Pupil Potential. Report Series No. 33.
INSTITUTION	Texas Univ., Austin. Research and Development Center for Teacher Education.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C. Bureau of Research.
BUREAU NO	BR-5-0249
PUB DATE	Sep 69
GRANT	OEG-6-10-108
NOTE	12p.
EDRS PRICE	EDRS Price MF-$0.25 HC-$0.70
DESCRIPTORS	*Academic Achievement, Elementary School Teachers, Secondary School Teachers, *Student Ability, *Student Evaluation, *Teacher Behavior

ABSTRACT

A study was conducted to examine the relative accuracy of elementary and secondary teachers in judging student capability. Fourteen eighth-grade teachers and 16 fifth-grade teachers were asked to rate special study children five times during the year on a behavior rating scale focusing on identification of under- and overachievement. From the pupils who met age (10, 14) and other criteria, four or five children were selected from each class for special study. Within each class were identified the boy and girl whose previous year's grade point average (GPA) rank in that class showed the greatest positive rank-order discrepancy with their California Test of Mental Maturity (CTMM) IQ rank in that class. The same plan was used to select the boy and girl exhibiting the greatest negative discrepancy. Thus a set of underachievers and overachievers were identified in each classroom. Teacher ratings were compared with the actual discrepancy scores between CTMM and GPA. Results demonstrated that teachers correctly identified more overachievers than underachievers. However, this was due to the fact that teachers identify more students as overachievers. Comparing the hit rate with the guess rate, it was concluded that teachers are not good judges of student capability. Secondary and elementary teachers appear equally poor judges. More research is needed to isolate the factors that prevent teachers from identifying children who are capable of performing at higher levels. (JS)

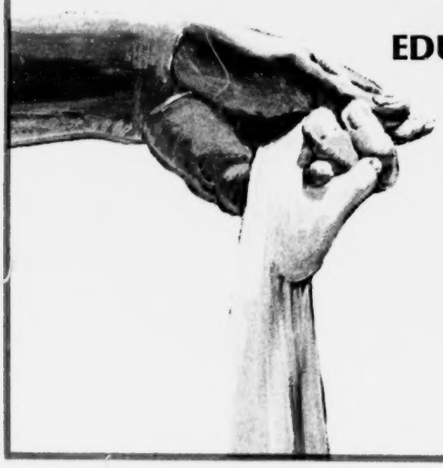
ED035589

TEACHER ASSESSMENT OF PUPIL POTENTIAL

Thomas L. Good
Donald L. Williams
Robert F. Peck
Linda M. Schmidt

Report Series No. 33

THE RESEARCH AND DEVELOPMENT CENTER FOR TEACHER EDUCATION



THE UNIVERSITY OF TEXAS A
AUSTIN / AUSTIN PUBLI
SCHOOLS / TEXAS EDU
CATION AGENCY

SP003447

BR 5 0 2 4 9
Report Series #2
PA-24
OE/BR

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

TEACHER ASSESSMENT OF PUPIL POTENTIAL

Thomas L. Good
Donald L. Williams
Robert F. Peck
Linda M. Schmidt

Report Series No. 33

September, 1969

The Research and Development Center for Teacher Education
The University of Texas at Austin

This research was supported by USOE Grant No. 6-10-108, The Research and Development Center for Teacher Education, Oliver H. Bown and Robert F. Peck, co-directors.

TEACHER ASSESSMENT OF PUPIL POTENTIAL

Thomas L. Good
Donald L. Williams
Robert F. Peck
Linda M. Schmidt

Introduction

Teachers evaluate student performance daily. At times evaluation is formal and permanent, a final mark. Frequently, the evaluation is spontaneous - a teacher responds quickly to student responses during class discussion. Teachers, by observing student behavior and by digesting cumulative file data, develop their personal view of student ability. This rating of student potential is a valuable step in determining future teaching behavior. The teacher may ask, "Does Johnny perform poorly because the material is too difficult for him presently or is his inadequate performance due to his indifference toward uninteresting and unchallenging assignments?"

Hadley (1954) examined the relationship between grades and measured achievement in 20 fourth-, fifth- and sixth-grade classrooms. He found the correlation between assigned marks and achievement (as measured by the California Achievement Test, Elementary Battery, Form AA) ranged from .20 to .94. Carter (1952) looked at the correlations of marks with achievement data in six classes and found for all boys and all girls the respective correlations to be .59 and .45. Both investigators questioned the accuracy of teacher marks. Carter contended that boys are graded lower

than their achievement merits and Hadley suggests that teachers tend to grade higher their most liked students and to grade lower their least liked students.

Teachers can, apparently, assign their own interpretation to student performance and these interpretations are not always isomorphic with actual accomplishment. They can identify students who do very well from students who do poorly, but can they determine differential student capability? Can they see the student who is capable of outstripping his present performance? If teachers are to maximize student learning, they must be able to match, with reasonable skill, student ability and task difficulty. The purpose of this paper is to examine the relative accuracy of elementary and secondary teachers in judging student capability.

Project Background

During the fall of 1968 a cooperative research program was initiated to help teachers to individualize their instruction to meet the needs of special study children. Teachers and university consultants, working in tandem, evaluated the effectiveness of various strategies for aiding children to cope more successfully with classroom life. Teachers had the opportunity to draw upon a comprehensive assessment battery, including achievement scores, intelligence scores, interests batteries and several projective instruments. In addition, teachers saw video tapes of classroom activities and met periodically with university researchers to discuss the special children.

As cooperating members of the research team, teachers made an intensive, continuing study of a few, selected children. Teachers were asked to rate the coping behavior of study children five times during the year. The Coping Behavior Rating Form (CBRF) is simply a system for describing how a child typically handles classroom problems. The teacher rates the child on 13 five-point scales such as, how competitive is he in school work, how does he feel about school and studies and how sociable is he.

Before the project began, teachers participated in a workshop, aimed at training them in the use of the CBRF. Special efforts were focused on getting teachers to recognize the meaning of each scale and to distinguish reliably between scale points. The desirability of using all scale points equally often was appropriately stressed. Examples representing each scale point on the 13 scales were presented to teachers for coding. After these ratings were collected, discussion ironed out rating discrepancies and ambiguous examples.

The CBRF manual contained the following instructions: "A child is rated at one of the five scale points for each question. The assumed group against which he is compared is that of all children of his age and sex in the national population. Thus, in a typical group of children, the ratings on any one question should be evenly spread along the entire scale, with some children high, some above average, some average, etc. In a large population, it would be assumed that about 20 percent of the children would fall at each of the five points on the scale."

Sample

The sample included 14 eighth-grade classes drawn from two junior high schools and 16 fifth-grade classes drawn from four elementary schools. From the pupils who met the age (10, 14) and SES criteria*, four or five children were selected from each class for special study. Within each class were identified the boy and girl whose previous year's GPA rank in that class, showed the greatest positive rank-order discrepancy with their IQ rank in that class. (IQ was measured by the California Test of Mental Maturity, CTMM). The same plan was used to select the boy and girl exhibiting

*The sample was drawn from that of a larger project, "Coping Styles and Achievement: a Cross-National Study of School Children," contract number: OE 5-85-063, which dealt with 10 and 14-year-old upper middle and upper lower class children in nine countries.

the greatest negative discrepancy. Thus a set of underachievers and overachievers were identified in each classroom.

Procedure

In September teachers were given the names of the special study children, two of whom were overachievers and two underachievers, selected in line with the criteria outlined in the sample description. The first teacher rating was made five weeks after school had commenced but before teachers were told the basis for child selection. Thus, the teacher ratings were uncontaminated with regard to the achievement classifications assigned their students, but the teachers presumably had access to cumulative file data in each school and could have gained the information themselves. Whether or not teachers used this information is unknown.

In an effort to look at the teacher as a judge of student ability, responses to scale VI of the coping rating form were examined with actual discrepancy scores for each child.

Scale Six was as follows:

VI. How good are the results he gets considering his ability?

(underachieves) 1 2 3 4 5 (Does his best)

Scale points 1 or 2 were considered as indicating that the child was underachieving and scale points 4 or 5 were interpreted as indicating that the child was overachieving. Scale point 3 seemed to represent a neutral classification and was not included in the following analyses. The teacher rating for each child was compared with his discrepancy score. The following definitions were employed in comparing teacher assessment of student ability and actual discrepancy scores:

Student capability: a discrepancy between measured classroom performance and measured student achievement.

Overachiever: positive discrepancy score between CTMM and grade point average.

Underachiever: negative discrepancy score between CTMM and grade point average.

Hits: Teacher assessment agreed with actual discrepancy score. That is, a teacher rating of 1 or 2 coupled with a negative discrepancy score or a 4 or 5 coupled with a positive discrepancy score.

HO: Teacher correctly identified an overachiever.

HU: Teacher correctly identified an underachiever.

Miss: Teacher assessments and discrepancy scores were in opposite directions. That is, a teacher ranking a 1 or 2 paired with a positive discrepancy score or a 4 or 5 paired with a negative discrepancy score.

MO: Teacher incorrectly ranked an overachiever.

MU: Teacher incorrectly ranked an underachiever.

Hypothesis

The investigators predicted that elementary teachers would be more accurate in their assessment of student ability than would be secondary teachers. Elementary teachers have much more exposure to their students than do secondary teachers. Elementary teachers spend their day with the same 30 students, while junior high teachers may work with 150 students a day. Further, a teacher with 30 students seems more likely to consult cumulative folders than would a teacher with a 150 students. Therefore, reason indicated that elementary teachers, having a greater opportunity to observe and to interact with their students, also would be more accurate in their assessment of student potential.

Results

Table 1 clearly shows that teachers were not accurate judges of student potential.

Table 1

Number of Hits and Misses, Between Teacher Rankings and Actual Discrepancy Scores

	HO	HO	H	MO	MU	M
Elementary	19	7	26	4	9	13
Secondary	11	3	14	4	15	19
Total	30	10	40	8	24	32

(H = hit, M = miss, O = overachiever, U = underachiever)

Teachers hit 40 cases and missed 32 cases. The hit rate is no better than chance. Upon further inspection of the data an interesting pattern was discovered. The distribution of Hits as shown in Table 1 shows that teachers hit 30 students who had positive discrepancy scores but hit only ten students who had negative discrepancy scores. Looking at the Misses it is seen that teachers missed 24 students who had negative discrepancy scores, but missed only eight students who had positive discrepancy scores. Apparently, teachers more sensitively identified over-achievers than underachievers.

Further inspection of the data however suggests that this conclusion is an artifact of the absolute number categorized in each cell. Table 2 shows the distribution of teacher response to scale VI.

Table 2

Distribution of Teacher Responses to Scale VI, "Guess Rate"

	Scale Points				
	1	2	3	4	5
Elementary	5	6	17	10	18
Secondary	2	5	6	12	14
Total	7	11	23	22	32

Table 2 indicates that teachers describe more children as over-achievers than underachievers ($p < .001$). The combined total of scale points 1 and 2 is 18 and the combined tallies of scale points 4 and 5 yields 54. (In actuality, 46 of the rated students had positive discrepancy scores and 49 had negative discrepancy scores). These figures make mandatory the comparison of hits and misses with the number of teacher guesses.

Table 3 displays the same hit and miss figures reported in Table 1.

Table 3

Distribution of Hit rate/Guess rate

	Hit overachiever	Hit underachiever
Elementary	19/28	7/11
Secondary	11/26	3/7
Total	30/54	10/18

However, in Table 3 these figures are compared with the number of students teachers identified as under and overachievers.

Looking at Table 3, one can see that teachers described children as underachievers on 18 occasions and were correct ten times. Teachers identified 54 children as overachievers and were correct 30 times. Teachers Hit more overachievers and Missed more underachievers because they identify more children as overachievers. Comparing the hit rate with the guess rate, teachers are evidently not better able to designate overachievers. Teachers, both secondary and elementary, appear equally poor judges of student performance. The data in Table 3 fail to confirm the initial hypothesis that elementary teachers would be better judges of student capability than would secondary teachers.

Discussion

The data are presented for the reader's examination, as little is known about the skills teachers possess for assessing student capability.

This was the first in a chain of experiments which, in sum, was designed to validate the use of teacher reports in assessing student behavior. Originally there were no serious doubts about elementary teachers - we expected them to assess student capability with modest precision. The major question was: How accurate are the assessments of secondary teachers who see students but a short time each day?

Results showed that neither elementary teachers nor secondary teachers were able to classify, reliably, student capability. Elementary teachers did tend to be more accurate and perhaps a larger sample would establish the fact that they are more aware of student capability. Although teachers do not accurately recognize student capability, this does not imply that teachers do not recognize the relative level of student performance within their class. They do. What they do not perceive is that some students are capable of performing better than at their present level.

Engaging in speculative proclamation, the investigators offer the following account as their interpretation of the data. Teachers do not expect much from their students, or at least teachers in de facto fashion sanction mediocre student performance. As only 18 students were identified by them as underachievers, teachers would seem to see students as doing okay. When teachers rate children as doing their best when, in fact, they are capable of doing more, the conclusion is that teachers are unable to identify many of their "academic" problems. So long as teachers perceive underachieving children as doing their best, these children will continue to go unchallenged and unnoticed.

The first step in providing a satisfying classroom life for children who underachieve is recognition of the problem. Recognition forces the teacher to examine her approach and say, "What can I do to help this student, what in the curriculum can I change?" To classify a child as an underachiever is a monumental step.

Such classification forces the teacher to take some action. She is acknowledging that the child can do better, but it is up to her to find the right key.

However, if the teacher denies or fails to perceive the student's potential, she has no dissonance, no problem. The child is doing his best, what you would expect of him; and this precludes any need to examine the adequacy of her approach. Quirk (1967) examined the locus of causality when students performed inadequately and when they performed admirably. Quirk found that when lessons went well and students were attentive, teachers assigned the credit to themselves; however, when lessons went poorly teachers shifted the blame to others. Perhaps the locus of causality needs to be examined in a variety of settings. In our study the possibility exists that teachers shifted the locus of causality by refusing to become aware of, or take seriously, information relevant to student potential.

One might contend that teachers were poor judges in this study because teacher training programs do not provide teachers with the skills necessary for analyzing the gap between performance and possible performance. If this is true, we might alternately claim that teachers cannot place children on a continuum between "underachievers" and "does his best," simply because they don't know how to do this. Irrespective of the precise determinant of teacher's inability to perceive student potential the resulting impact is the same: a loss of human potential.

More research is needed to pinpoint the accuracy of teachers in judging student capability. If this sample is representative of teachers in general, the picture is bleak. Many children capable of performing at higher levels go unnoticed. Apparently, such potential is left on the vine to wither. They are labeled as doing their best, no problem, Doubtlessly, failure to identify the potential of these children leads to instructional mediocrity. But perhaps the greatest loss is when such children come themselves to view their present performance as adequate.

Summary

Fourteen eighth-grade teachers and 16 fifth-grade teachers were asked to rate special study children on a behavior rating scale. Special attention was focused on scale VI, "How good are the results he gets, considering his ability?" Teachers responded to this question by marking a scale point from 1 to 5 (from "underachiever" to "does his best"). Teacher responses 1 and 2 were interpreted as an identification of an underachiever and a 4 or 5 was interpreted as a teacher identification of an over-achiever.

From the pupils who met the age (10, 14) and SES criteria (UL, UM) four or five children were selected from each class for special study. Within each class were identified the boy and girl whose previous year's GPA rank in that class, showed the greatest positive rank-order discrepancy with their CTMM IQ rank in that class. The same plan was used to select the boy and girl exhibiting the greatest negative discrepancy. Thus a set of underachievers and overachievers were identified in each classroom.

Teacher ratings were compared with the actual discrepancy scores between CTMM and grade point average. Results demonstrated that teachers correctly identified more overachievers than under-achievers. However, this was due to the fact that teachers identify more students as overachievers. Comparing the hit rate with the guess rate, an apparent conclusion is that teachers are not good judges of student capability.

Teachers do not perceive student capability, perhaps because they have not been trained to do so or because they deliberately ignore cumulative file data. The consequence is that teachers accept minimum student performance, perceiving children we would call underachievers, as doing okay. Hence, these children go unnoticed, and are not identified as children who can exceed their present classroom performance. Identification of an under-achiever is the first step in helping the child; more research is needed to isolate the factors that prevent teachers from identifying children who are capable of performing at higher levels.

References

Carter, Robert S. How Invalid Are Marks Assigned by Teachers? Journal of Educational Psychology, 43, (1952), pp. 218-28.

Hadley, S. Trevor. "A School Mark -- Fact or Fancy?," Educational Administration and Supervision, 40: 305-312, (May, 1954).

Quirk, Thomas J. An Experimental Investigation of the Teacher's Attribution of the Locus of Causality of Student Performance. Dissertation Abstracts, 28, 2565 - A, 1967/1968. (Order no. 67-17, 487.)